PCT/EP 0 3 / 0 9 8 8 2

BUNDESREPUBLIK DEUTSCHLAND

06. 10. 2003
10/525872

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 41 207.3

Anmeldetag: 05. September 2002

Anmelder/Inhaber: CELL CENTER COLOGNE GmbH, Köln/DE

Bezeichnung: Immunogenic MUCI glycopeptides

IPC: C 07 K 14/435

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 25. September 2003
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Eberl

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

A 9161
02/00
EDV-L

Abstract:

The present invention is directed to MUC1 peptide fragments and to methods of producing those peptide fragments. The invention is further directed to an ex vivo method of producing a population of autologous antigen presenting cells (APC's) and of producing genetically engineered APC's, which are capable of inducing effective immune responses against MUC1. The invention also relates to APC's, which are obtainable by these methods as well as the use of the above mentioned fragments and APC's in a therapeutic composition for the treatment of breast cancer or other MUC1-positive carcinomas including colorectal, pancreatic and gastric carcinomas.

P14655

Immunogenic MUC1 glycopeptides

The present invention is directed to MUC1 peptide fragments and to methods of producing those peptide fragments. The invention is further directed to an ex vivo-method of producing a population of autologous antigen presenting cells (APC's) and of producing genetically engineered APC's, which are capable of inducing effective immune responses against MUC1. The invention also relates to APC's, which are obtainable by these methods as well as the use of the above mentioned fragments and APC's in a therapeutic composition for the treatment of breast cancer or other MUC1-positive carcinomas including colorectal, pancreatic and gastric carcinomas.

MUC1 is overexpressed in breast cancer and by many other carcinomas and the tumor-associated glycoform of the mucin is known to expose multiple peptide epitopes within its repeat domain. These immunogenic peptide epitopes make MUC1 a promising tumor antigen with diagnostic as well as therapeutic potential in the treatment of cancer.

The development of effective vaccine and immunotherapies for human cancers and infectious agents often is dependent on the generation of protective immune responses to specific domains of membrane proteins. The tandem repeat (TR) domain of the breast, pancreatic, and ovarian tumor antigen, human mucin MUC1 (Barnd et al., PNAS USA, 86: 7159-7163, 1989; Jerome et al., Cancer Res., 51: 2908-2916, 1991), the principal neutralizing domain of HIV-1 (Javaherian et al., PNAS USA, 86: 6768-6772, 1989; Javaherian et al., Science, 250: 1590-1593, 1990) and the proline rich neutralization domain of the feline leukemia virus external surface unit protein (gp-70) (Nunberg et al., PNAS, 81: 3675-3679, 1984; Elder et al., J. Virol., 61: 8-15, 1987; Strouss et al., J. Virol., 61: 3410-3415, 1987; Nick et al., J. Gen. Virol., 71: 77-83, 1990) are examples hereof.

Regarding MUC1, humoral and cellular responses have been demonstrated in cancer patients (1;2), but also in pregnant woman (3) and healthy individuals (4). Although these natural responses are usually insufficient to fight the progress of cancer, MUC1-derived peptides or

glycopeptides are used currently in clinical trials to trigger therapeutically and prophylactically immune reactions in humans (5,6).

There is growing evidence that triggering of efficient humoral and CTL responses to MUC1 needs the activation of specific T helper cell clones, which is induced by MHC class II-presented antigen fragments. The generation of MHC class II-restricted peptide epitopes by antigen presenting cells (APCs) like dendritic cells (DCs) follows a multistep process starting with endocytosis, followed by the processing in late endosomal compartments and resulting in the binding of proteolytic peptide fragments to MHC class II proteins and their transport to the cell surface. While many aspects of this complex process have been elucidated there is currently little evidence on the processing and MHC class II presentation of glycosylated antigens, in particular of the highly O-glycosylated mucin antigens. To enable the design of efficient tumor vaccines on the basis of MUC1 knowledge on how DCs or other APCs deal with O-glycosylated peptides is of importance. One particular question in this context refers to the fate of complex O-linked glycans during processing, since efficient peptide fragmentation may afford complete or partial removal of sugars prior to proteolysis. O-linked glycans could also direct the processing with respect to the accessibility of cleavage sites and hence restrict the pattern of peptide fragments on the one hand, while they enrich the pattern of epitopes on the other. The inventors have recently shown by immunological methods that MUC1 core-glycans are not removed during processing and that glycopeptides presented by MHC class II are able to activate T cell hybridomas (unpublished).

Although several of the cathepsins have been identified as components of the processing machinery (8,9), it is currently not known which enzyme(s) are involved in the processing of MUC1 and at which sites within the repeat domain they actually cleave the protein. Expectedly, there are multiple cleavage sites and only subfractions of the generated peptide fragments should fulfill the requirements for binding to MHC class II molecules. Thus, it could be a valuable approach for the design of new, anti-cancer vaccines to identify new immungenic MUC1 fragments capable of binding to MHC class II molecules.

Therefore, it is an object of the present invention to provide novel immungenic MUC1 peptide fragments, which can be used for immunization in humans.

It is a further object of the present invention to provide a method of producing an immunogenic MUC1 peptide, which allows the originally contained glycosylation pattern to be conserved during the production process.

These objects are solved by the subject-matters of the independent claims. Preferred embodiments are set forth in the dependent claims.

Surprizingly, the inventor has found out that cathepsin L or a closely related enzyme shows a very restricted fragmentation pattern during mouse DC processing with only one preferred cleavage site per MUC1 repeat. The cleavage specificity and specific inhibition of the protease were in agreement with the assumption of the inventor that cathepsin L or a closely related enzyme (cathepsins B or S) were involved in this highly specific cleavage.

The experimental set-up used biotinylated and non-tagged, bead-coated synthetic glycopeptides comprising one or more repeat units of MUC1 with single or multiple O-linked core-type glycans. Exogenously administered MUC1 peptide fragments were rapidly taken up by mouse dendritic cells (DCs) and a large proportion was processed in late endosomal compartments during 4h. MUC1 repeat peptide derived proteolytic fragments that were identified and sequenced show that the glycans are not removed during antigen processing and that the presence of carbohydrates affects the cleavage sites yielding a different repertoire of cleaved peptides.

Surprizingly, the proteolytic products suggest a highly specific processing of the repeat peptide with one preferential cleavage site at the Thr-Ser peptide bond. While human cathepsin D was unable to cleave the MUC1 repeat peptide *in vitro*, human cathepsin L digestion resulted in specific hydrolysis of the Thr-Ser peptide bond. Since MUC1 sequences contain a VTSA motif in every repeat unit, the generated fragments start with the amino acid sequence SAP at their N-terminus.

Information on the structure of processed MUC1 glycopeptides is of utmost importance for the design of tumor vaccines. Intact O-glycosylation on processed MUC1 repeat peptide contributes to a greater variety of the MHC class II-restricted helper T cell responses, thereby enhancing an overall anti-tumor response.

According to the invention, a MUC1 peptide fragment of the tandem repeat domain of human MUC1 is provided, starting with the amino acid sequence SAP and comprizing at least 9 amino acids. The amino acid sequence of human MUC1 is already known and can be found, for example, in the SWISS PROT database. The MUC1 protein contains varying numbers of amino acids due to a length polymorphism resulting from individually variable repeat numbers, and, in the moment, at least 9 isoforms are known (1/A, 2/B, 3/C, 4/D, 5/SEC, 6/X, 7/Y, 8/Z and 9/S, which are produced by alternative splicing).

In this invention, specific fragments of MUC1 are contemplated, which are derived from the a synthetic or natural MUC1 sequence, which has been cleaved enzymatically at the VTSA motif contained in all MUC1 sequences (or was chemically synthesized in case of synthetic fragments). The fragments of the present invention thus can be obtained by cleavage of MUC1 sequences with cathepsin-L. Irrespective of the starting amino acid position in the repeat sequence (TAP, AHG, GST) and of the length of the peptides (20mer, 21mer, 25mer, 100mer), cathepsin L cleaves specifically between Thr-Ser in the VTSA motif of the repeat peptide, thereby resulting in the inventive peptide fragments. It is an essential feature of the present invention that all fragments start with the amino acid sequence SAP at their N-terminus.

As mentioned above, the MUC1 peptide fragment of the present invention is not limited in its length, and may, for example, comprise up to 100 amino acids or even more. However, amino acids are preferred, wherein the MUC1 fragment is in the range of 10 to 25 amino acids.

According to a further embodiment, the invention provides specific MUC1 peptide fragments which comprise the amino acids of SEQ ID NO 1-4 or variants thereof, wherein said variants comprise one or more insertions, substitutions and/or deletions as compared to the sequence of SEQ ID No. 1-4, and wherein the biological activity is substantially equal to the activity of the peptide comprising the unmodified amino acid sequence of SEQ ID No. 1-4, and provided that said variants start with the amino acid sequence SAP and comprise at least 9 amino acids.

In this context, the present invention provides the following peptides:

SAPDTRPAPGSTAPPAHGVT (SEQ ID NO 1)

SAPESRPAPGSTAPAAHGVT (SEQ ID NO 2)

SAPESRPAPGSTAPPAHGVT (SEQ ID NO 3)

SAPDTRPAPGSTAPAAHGVT (SEQ ID NO 4)

The arrow indicates that the present invention also encompasses variants of the above mentioned amino acid sequences, which are reduced by one or more amino acids starting from the C-terminus, under the proviso that the variants at least comprise the 9 N-terminal amino acids of the above indicated sequences (bold printed).

The term biological activity as used herein is related to the immunogenic function of the inventive amino acid sequences. As mentioned above, MUC1 is naturally overexpressed in various cancers, like breast cancer and other adenocarcinomas, and therefore, it is an important target for immune based anti-cancer therapy. Thus, the MUC1 fragments as disclosed hereinbefore are contemplated as long as they are capable of inducing an immungenic reaction in mammals, preferably humans, in order to initiate/promote an attack of the patient's immune system against the respective cancer.

The amino acid sequences of the present invention also encompass all sequences differing from the herein disclosed sequences by amino acid insertions, deletions, and substitutions.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, i.e., conservative amino acid replacements. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

"Insertions" or "deletions" are typically in the range of about 1 to 3 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or

substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity. This does not require more than routine experiments for the skilled artisan. In case of MUC1 repeats three positions are known to exhibit a sequence polymorphism in the population (Engelmann et al., 2001, J. Biol. Chem. 276, 27764 – 27769; PCT patent application PCT/DE00/00440, the disclosure thereof being incorporated in its entirety in this application).

The present invention is further directed to a nucleic acid, which encodes one of the above mentioned MUC1 peptide fragments.

The term "nucleic acid sequence" refers to a heteropolymer of nucleotides or the sequence of these nucleotides. The terms "nucleic acid" and "polynucleotide" are used interchangeably herein to refer to a heteropolymer of nucleotides.

The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes to the complement of the disclosed nucleotide sequences under moderately stringent or stringent hybridization conditions; a polynucleotide which is an allelic variant of any polynucleotide recited above; a polynucleotide which encodes a species homologue of any of the herein disclosed proteins; or a polynucleotide that encodes a polypeptide comprising an additional specific domain or truncation of the disclosed proteins.

Stringency of hybridization, as used herein, refers to conditions under which polynucleotide duplexes are stable. As known to those of skill in the art, the stability of duplex is a function of sodium ion concentration and temperature (see, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd Ed. (Cold Spring Harbor Laboratory, (1989)). Stringency levels used to hybridize can be readily varied by those of skill in the art.

Low stringency hybridization refers to conditions equivalent to hybridization in 10% formamide, 5 x Denhart's solution, 6 x SSPE, 0.2% SDS at 42°C, followed by washing in 1 x SSPE, 0.2% SDS, at 50°C Denhart's solution and SSPE are well known to those of skill in the art as are other suitable hybridization buffers.

Moderately stringent hybridization refers to conditions that permit DNA to bind a complementary nucleic acid that has about 60% identity, preferably about 75% identity, more

preferably about 85% identity to the DNA; with greater than about 90% identity to said DNA being especially preferred. Preferably, moderately stringent conditions are conditions equivalent to hybridization in 50% formamide, 5 x Denhart's solution, 5 x SSPE, 0.2% SDS at 42°C, followed by washing in 0.2 x SSPE, 0.2% SDS, at 65°C.

High stringency hybridization refers to conditions that permit hybridization of only those nucleic acid sequences that form stable duplex in 0.018M NaCl at 65°C. (i.e., if a duplex is not stable in 0.018M NaCl at 65 °C, it will not be stable under high stringency conditions, as contemplated herein).

Further, nucleic acid hybridization techniques can be used to identify and obtain a nucleic acid within the scope of the invention. Briefly, any nucleic acid having some homology to a sequence set forth in this invention, or fragment thereof, can be used as a probe to identify a similar nucleic acid by hybridization under conditions of moderate to high stringency. Such similar nucleic acid then can be isolated, sequenced, and analyzed to determine whether they are within the scope of the invention as described herein.

According to a preferred embodiment, the MUC1 peptide fragments of the present invention are O-glycosylated at one or more of the threonines or serines contained in the sequence. Preferably, the MUC1 peptide fragments of one of SEQ ID No. 1-4 are glycosylated at Thr 5 and/or 12. However, also all other serines or threonins in the respective sequences may be glycosylated. A preferred glycan used herein is GalNAc or further complex glycans, which are derived therefrom.

According to a further aspect, the present invention provides a method of producing the inventive MUC1 peptide fragments, comprising the following steps:

a) providing a peptide comprising the tandem repeat domain of MUC1 or a part thereof, which part at least contains one repeating unit of said tandem repeat domain of MUC1,

b) contacting the peptide of a) with an effective amount of cathepsin-L, thereby cleaving the peptide, and

c) isolating the fragments produced in b).

Preferably, the peptide provided in a) is a MUC1 protein showing a natural glycosylation pattern. As mentioned above, it was surprisingly found out by the present inventor that a

cathepsin-L cleavage as performed in step b), leaves the glycosylation pattern of the MUC1 protein, provided in a), intact. Intact O-glycosylation on processed MUC1 repeat peptides in turn contributes to a greater variety of the MHC class II-restricted helper T cell responses, thereby enhancing an overall anti-tumor response in patients. Thus, the inventive method leads to a MUC1 peptide fragment, which can be easily processed by the patient's APC's, for example dendritic cells, by the MHC class II pathway, and will be presented with an intact glycosylation pattern leading to an enhanced immune response of helper T-cells. In this context it should be noted that there is no restriction regarding the glycosylation pattern, however, threonin glycosylated at the cleavage site leads to a Thr-Ser bonding, which is stable to cathepsin L proteolysis.

Of course, the above mentioned method is not the only one which leads to the inventive peptide fragments, whether glycosylated or not. It is also possible to chemically synthesize those fragments thereby providing, for example, a desired glycosylation pattern. To synthesize glycopeptides, glycosylamino acid building blocks are required which already contain the oligosaccharide chain and threonine or serine. The syntheses of these building blocks have been described (Mathieux, N., Paulsen, H., Meldal, M., Bock, K. J.Chem.Soc., Perkin Trans. *1*: 2359-2368, 1997). The multiple column solid phase synthesis can be carried out in a semi-manual 20-column multiple synthesizer, and Wang resin can be selected as support material. The Wang resin (2,5 g) can for example be placed in a glass reactor, swelled in dichloromethane (15 cm^3, 10 min.) and washed. A mixture of Fmoc-Ala-OH (3,40 mmol), 1-(mesitylenesulfonyl)-3-nitro-1,2,4-triazole (3,40 mmol) and methylimidazole (3,40 mmol) in dichloromethane (15 cm^3) was added. After 2 h, the resin can be washed and the unchanged amino groups can be acetylated with Ac_2O/DMF (1:1; 15 cm^3). The derivatized resin is then packed for the glycopeptide synthesis in the 20 columns of the synthesizer. The reaction and washing solvent can be DMF, the Fmoc deprotections were performed by treatment with piperidine (20 %) in DMF (20 min.). The amino acids are coupled as Fmoc amino acid Pfp ester with Dhbt-OH (3 mol equiv.). The Gal(1→3)GalNAc-containing building block are coupled with TBTU and N-ethyldiisopropylamine (1,5 mol equiv.). After 20 h reaction time the synthesis cycle is repeated to complete the assembly of each glycopeptide. After removal of the last Fmoc groups, the resins are washed, dried, treated with 95 % aq TFA (2 cm^3, 2 h), and filtered off. Then, the compounds is treated with catalytic amounts of 1 % CH_3ONa in methanol at pH 8,5 to remove the acetylic groups of the saccharide part, and purified by

preparative RP-HPLC. The pure O-glycopeptides are obtained in yields of 16-57 % after lyophilization.

Preferably, glycopeptides are formed containing O-linked GalNAc or elongated complex glycans at one or several of the threonine or serine residues.

Preferably, the peptide provided in step a) is represented by natural MUC1 derived from human milk fat membranes (see Müller et al., 1997, J. Biol. Chem. 272, 24780 – 24793), from tumor ascites (Beatty et al., 2001, Clin. Cancer Res. 7, 781 – 787) or from human breast carcinoma cell lines (Müller et al., 2002, J. Biol. Chem. 277, 26103 – 26112) or is represented by SEQ ID NO 5, 6, 9 or 10.

Furthermore, the inventive amino acids of the peptide provided in step a) of the above method of producing the inventive MUC1 peptide fragments are O-glycosylated, however, provided that the peptide is not glycosylated at the cleaving site of cathepsin-L. Preferably, one or more of the threonines or serines of the inventive peptide fragment isolated in c) are O-glycosylated.

According to a further aspect, a MUC1 peptide fragment is provided, which is obtainable by the above mentioned methods.

The invention is further directed to an ex vivo-method of producing a population of autologous antigen presenting cells (APC's), which is capable of inducing effective immune responses against MUC1, comprising the steps of

a) contacting the autologous APC's from tumor patients with an effective amount of an inventive MUC1 peptide fragment under conditions which allow endocytosis, processing and MHC class II presentation of the peptide fragments by said APC's, and

b) isolating said MUC1 peptide presenting APC's for the purpose of immunotherapeutic application in patients.

Preferably, the MUC1 peptide fragments in a) are bound to coated ferric oxide beads. However, it is noted that all other known beads can be used for the purpose of the above mentioned method. Generally, all beads can be used, which are not larger than approx. 1-2 µm in size and allow a covalent coupling of antibodies and lectines.

Furthermore, an ex vivo-method of producing genetically engineered APC's is provided, which are capable of inducing effective immune responses against MUC1, comprising the steps of

a) providing a nucleic acid, which codes for one of the inventive MUC1 peptide fragments;

b) transfecting the APC's with said nucleic acid, and

c) selecting APC's, which present said MUC1 peptide fragments in a MHC II restricted manner.

According to a preferred embodiment, the nucleic acid in step a) is provided in an expression vector. This expression vector preferably comprises one or more regulatory sequences. The term "expression vector" generally refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vector can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

According to a further aspect of the invention, an APC is provided, which is obtainable by one of the aforementioned methods. Preferably, this APC is a dendritic cell or a B cell.

Furthermore, the present invention provides a therapeutic composition, comprising the inventive MUC 1 peptide fragment or the inventive APC's and a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to the ingredient and the carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The therapeutic composition may

further contain other agents which either enhance the activity or use in treatment. Such additional factors and/or agents may be included in the therapeutic composition to produce a synergistic effect or to minimize side-effects.

Techniques for formulation and administration of the compounds of the present application may be found in "Remington's Pharmaceutical Sciences", Mack Publishing Co., Easton, PA, latest edition.

The compositions contain a therapeutically effective dose of the respective ingredient. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of such conditions, specifically in an induction of an immune response in the patient. Suitable routes of administration may, for example, include parenteral delivery, including intramuscular and subcutaneous injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal injections. Intravenous administration to the patient is preferred.

A typical composition for intravenous infusion can be made up to contain 250 ml of sterile Ringer's solution, and 10 mg of the ingredient. See Remington's Pharmaceutical Science (15th Ed., Mack Publishing Company, Easton, Ps., 1980). Preferably, the therapeutic composition of the present invention is a vaccine.

As mentioned above, this vaccine finds application for use in the treatment of breast cancer or other MUC1-positive carcinomas including colorectal, pancreatic and gastric carcinomas.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The invention is now further illustrated by the accompanying drawings, which are showing the following:

Fig. 1

MUC1 glycopeptide processing by mouse dendritic cells

A mixture of biotinylated glycopeptides H1 to H3 (AHGVTSAPDTRPAPGSTAPPA; SEQ ID NO 5) and H4 to H6 (AHGVTSAPESRPAPGSTAPAA; SEQ ID NO 6) corresponding to a partial sequence of MUC1 tandem repeats and glycosylated with GalNAc at Thr5 (H1, H4), Thr10/Ser10 (H2, H5) or Thr17 (H3, H6) was used for pulsing of mouse dendritic cells D2.4. Processing products were affinity-isolated from cellular fractions or from culture supernatants, reduced with dithiothreitol to cleave the label, and analysed by reflectron MALDI mass spectrometry in the positive ion mode.

A, centrifuged cellular fraction;

B, magnetically separated cellular fraction;

C, culture supernatant

The major signals at m/z 2249.0 (H1 to H3) and 2223.0 (H4 to H6) correspond to the precursor glycopeptides, the signals at m/z 1695.7 (P1) and 1669.7 (P2) to the SAP16 fragments (P1 derived from H1 to H3; P2 from H4 to H6).

Fig. 2

Peptide sequencing of processing products P1 and P2 by LC-MS/MS analysis on a Qtof2 electrospray mass spectrometer

(Glyco)peptides were separated by nanoflow liquid chromatography on a reversed-phase microcapillary column and analysed online by electrospray mass spectrometry in the positive ion mode. B ion and y ion fragment series from the N-terminal and C-terminal sequences of the major peptide products (A, P1 at m/z 1695; B, P2 at m/z 1669) were used to confirm the sequence of SAP16 peptides (refer to C).

Fig. 3

In vitro proteolysis of MUC1 glycopeptides by human cathepsin L

MUC1 glycopeptides (10 µg) were treated for 3h with 2 to 5 munits of cathepsin L in 0.1M sodium acetate, pH 5.5, containing 1 mM EDTA, and 1 mM DTT. Reflectron MALDI mass spectra were recorded in the positive ion mode using α-cyano-4-hydroxy cinnamic acid as matrix.

A, Biotinylated A3 glycopeptide (m/z 2784.7), A3 glycopeptide after reductive cleavage of the label (m/z 2412.6), SAP16 fragment (m/z 1858.3).

B, Products of non-tagged A3 glycopeptide: SAP16 fragment (m/z 1858.0), APD15 (m/z 1771.0), TSA17 (m/z 1959.0).

Fig. 4

Cathepsin L-like activity in low-density endosomes from mouse dendritic cells cleaves MUC1 repeats at Thr-Ser

Low-density endosomes in mouse dendritic cells were separated from lysosomes and plasma membranes by density gradient centrifugation in percoll/sucrose (30 ml). Fractions (1 ml) were tested for the presence of marker proteins/enzymes (β-hexosaminidase, H2 antigen) and for proteolytic activity. TAP25 peptide was used as substrate (20h; 37 °C, pH 5.5, 1 mM EDTA, 1 mM DTT).

Materials and Methods

Isolation and cultivation of dendritic cells

Immortalized dendritic cells (clone D2.4) from C57BL/6 mice was grown in DMEM supplemented with 10% FCS, L-glutamine, 0.1% 2-mercaptoethanol, and antibiotics at 37°C and 5% CO_2 (9).

Generation of synthetic (biotinylated) MUC1 glycopeptides and their conjugation to magnetic ferric oxide beads

Glycopeptides H1 to H6 corresponding to MUC1 tandem repeat peptides based on the AHG21 sequences AHGVTSAPDTRPAPGSTAPPA (H1 to H3) and AHGVTSAPESRPAPGSTAPAA (H4 to H6) and carrying GalNAc at Thr5, Thr10, or Thr17 were chemically synthesized according to previously published protocols (10) and isolated successively on preparative and analytical reversed-phase columns on an HPLC workstation (System Gold, Beckman, München, Germany). The same holds true for the glycopeptide series H11 to H13 (GalNAc substituted) and A3 (substituted with Galβ1-3GalNAc at Thr17), which are based on the same peptide sequence as H1 to H3. The 100mer peptide corresponding to five repeats of the MUC1 domain and starting with the HGV motif was synthesized by a local facility (University of Pittsburgh) and *in vitro* glycosylated with GalNAc using purified polypeptide GalNAc-transferases-T1 and –T2 (kindly provided by Dr.

Henrik Clausen, School of Dentistry, University of Copenhagen, Denmark) under conditions described previously (11,12).

Glycopeptides H1 to H6 (100 µg each) were biotinylated with [2-(Biotinamido)ethylamido]-3,3'-dithiopropionic acid N-hydroxysuccinimide ester (100 mM in DMSO, 100 µl) at 50°C over a period of 48h. After evaporation of the solvent by vacuum centrifugation the biotinylated products were separated from non-tagged glycopeptides and excessive reagent by reversed-phase chromatography on a PLRP-S column.

Anti-MUC1 dynabeads were prepared by covalent coupling of 50 µg B27.29 monoclonal antibody to tosylated M-280 beads (Dynal, Hamburg, Germany) in 0.1 M borate buffer, pH 9.5 (200 µl) for 48h at ambient temperature. Lectin-coated dynabeads were prepared similarly by conjugation of 50 µg Helix pomatia agglutinin to M-280 beads. Antibody- and lectin-coated beads (10^8) were complexed with (biotinylated) glycopeptides (50 µg) by incubation in 250 µl AIMV medium under rolling for 2h at ambient temperature.

Antigen pulse of dendritic cells and isolation of peptide fragments

Mouse dendritic cells D2.4 (10^7 cells/ml) were transferred into a 15 ml Falcon tube, suspended in AIMV medium and preincubated for 1h at 37°C (5% CO_2). Antigens were added as a mixture of biotinylated glycopeptides H1 to H6 (50 µg) after conjugation to anti-MUC1 antibody- and lectin (HPA)-coated dynabeads (each at 5 x 10^7 beads / ml final concentration). The 1 ml suspension was incubated with occasional shaking at 37°C (5% CO_2) for a total time period of 4h. After pulsing the cells were separated from the medium in two ways: an orthogonal magnetic separation of bead-loaded cells and free dynabeads and a centrifugation of the remaining cells (180 g, 5 min), which had not been removed in the first step due to a lower antigen/bead load. Both cell fractions were washed several times in phosphate (4 mM), NaCl (153 mM), pH 7.2, while the cell-free supernatant was re-centrifuged at 3000 g (5 min, 4°C). The cell fractions were treated on ice for 15 min with 100 µl 1% NP40, 10 mM Tris-HCl, 150 mM NaCl, pH 8.0 containing a cocktail of protease inhibitors (Sigma P8340, München, Germany) followed by ultrasonication for 2 min.

In experiments where biotinylated antigens had been used the cell extracts were diluted twofold with PBS and incubated with 2 x 10^8 streptavidine-coated dynabeads M-270 for 30

min at 37°C and another 30 min period with rolling at ambient temperature. After magnetic separation and washing of the beads for three times the biotinylated glycopeptides were cleaved-off by reduction with 10 mM dithiothreitol at 56°C (30 min), dried by vacuum centrifugation and taken up in 0.1% aqueous trifluoroacetic acid. Since both cell fractions contained considerable proportions of antibody- and lectin-coated beads, also non-tagged peptide and glycopeptide fragments were isolated during this separation step. Considerable amounts of non-tagged MUC1 glycopeptides were also demonstrated to bind to streptavidine-coated dynabeads via undefined mechanisms and to elute during heating under reducing conditions. Peptides and glycopeptides contained in the cell-free supernatant were affinity-isolated by using mixtures of streptavidine-, anti-MUC1 antibody-, and lectin-coated beads or by applying reversed-phase chromatography on 50 µl columns of Poros C18. After loading of the reversed-phase column with 500 µl of supernatant the sample was desalted by washing with 0.1% aqueous trifluoroacetic acid and eluted with 80% acetonitril in 0.1% aqueous trifluoroacetic acid.

Mass spectrometric analyses

MALDI mass spectrometry: The peptide and glycopeptide samples (20 µl) contained in 0.1% aqueous trifluoroacetic acid or in mixtures with acetonitril were applied to the stainless steel target by mixing a 1 µl aliquot with the same volume of matrix (saturated solution of α-cyano-4-hydroxycinnamic acid in acetonitril / 0:1% TFA, 2:1). Mass spectrometric analysis was performed on a Bruker-Reflex IV instrument (Bruker-Daltonic, Bremen, Germany) by positive ion detection in the reflectron mode. Ionization of co-cristallized analytes was induced with a pulsed nitrogen laser beam (337 nm) and the ions were accelerated in a field of 20 kV and reflected at 23 kV (12, 13).

Nanoflow liquid chromatography with on-line ESI mass spectrometry: LC/MS data were acquired on a Q-Tof II quadrupole-time of flight mass spectrometer (Micromass, Manchester, UK) equipped with a Z spray source. Samples were introduced using the Ultimate nano-LC system (LC Packings, Amsterdam, Netherlands) equipped with the Famos autosampler and the Switchos column switching module. The column setup comprised a 0,3 mm x 1 mm trap column and a 0,075 x 150 mm analytical column, both packed with 3 µm PepMap C18 (LC Packings, Amsterdam, Netherlands). Samples were diluted 1:10 in 0,1 % TFA. 10 µl were injected onto the trap column and desalted for 3 min using 0,1 % TFA and a flow rate of 30

µl/min. The 10 port valve switched the trap column into the analytical flowpath and the peptides were eluted onto the analytical column using a gradient of 5 % ACN in 0,1 % formic acid to 40 % ACN in formic acid over 20 min and a column flow rate of approximately 200 nl/min, resulting from a 1:1000 split of the 200 µl/min flow delivered by the pump. The ESI interface comprised a metal coated PicoTip spray emitter (New Objective, MA) mounted onto the PicoTip holder assembly (New Objective). Stable nanospray was established by the application of 2,5 – 3,0 kV to the distal end of the PicoTip and a nitrogen counter flow rate of approximately 40 L/min. Data dependent acquisition of MS and MS/MS spectra was controlled by the masslynx software. Survey scans of 1 sec covered the range from m/z 400 to m/z 1200. Doubly and triply charged ions rising above a given threshold were selected for MS/MS experiments. In MS/MS mode the mass range from m/z 40 to m/z 1400 was scanned in 1 sec and 10 scans were added up for each experiment. Micromass formated peaklists were generated from the raw data using the Proteinlynx software module.

Confocal Laser Scanning Microscopy and fluorescence-activated cell sorting

Antigen uptake was quantitated by flow cytometric analysis using a Becton Dickinson FACScalibur according to a previously published protokol (14). Prior to microscopic inspection DCs were fixed with 2% formaldehyde, and permeabilized with 0.1% saponin. Following staining with anti-MUC1 antibody (B27.29, Biomira, Edmonton, Canada), biotinylated secondary anti-mouse Ig (Dako, Hamburg, Germany) and FITC-labelled streptavidine (Sigma), the cells were fixed a second time with 1% paraformaldehyde, the chambers of the slides were removed, and the slides were mounted for the analysis by confocal laser scanning microscopy on a Leica DM IRE2 (14).

In vitro proteolysis of MUC1 (glyco)peptides with human cathepsins and low-density endosomal fractions from mouse dendritic cells

Human cathepsin L and D were purchased from Sigma (München, Germany) and solubilized in 0.1 M sodium acetate buffer, pH 5.5, containing 1 mM EDTA (cathepsin D) and 1 mM dithiotreitol (cathepsin L). 2 – 5 munits of enzyme(s) were added to 10 µg of (glyco)peptide substrates (corresponding to approx. 5 nmol of repeat units) in a total volume of 20 µl digestion buffer (see above). The reaction mixtures were incubated at 37°C and 2 µl were

withdrawn after 3h or 24h and diluted 10 to 20fold in 0.1% aqueous TFA prior to MALDI mass spectrometry.

Mouse dendritic cells (10^8) were homogenized by fine-needle aspiration on ice using 1 ml of 0.3 M sucrose, 0.01 M Hepes as buffer (without protease inhibitors). After dilution to 7 ml and centrifugation at 850 g for 10 min to remove intact cells and nuclei, 6 ml of the supernatant were centrifuged over 24 ml of 35% Percoll with 0.3 M sucrose, 0.01 M Hepes for 105 min at 20.000 rpm in a centrifuge (model J2-21 M/E, rotor: JA-20, Beckman instruments, München, Germany) (15). The gradient was fractionated by gravity siphon (30 x 1 ml) and each fraction was analysed after sonication for the presence of MHC class II molecules by enzyme immunoassay with anti-H2 antibody, β-hexosaminidase activity (15) and cathepsin L related proteolytic activity using TAP25 peptide as substrate (5 µg). The samples were incubated for 24h at 37°C, diluted 20fold in aqueous TFA and analysed by MALDI mass spectrometry. For specific inhibition of cathepsin L activity, the corresponding fractions were mixed with 1 µM Z-Leu-Leu-Leu-fluoromethyl ketone (Sigma).

Processing of MUC1 glycopeptides by mouse dendritic cells

In a series of experiments a variety of application modes for the pulsing DCs with antigens were tested by using confocal laser microscopy and fluorescence activated cell sorting in the kinetic analysis of Cy3-labelled MUC1 peptide (100mer). According to quantitative FACScan data the proportion of fluorescently labelled cells increased with time up to 2h, but even more with the mode of antigen application as free peptide, as peptide-antibody complex, or as conjugates to paramagnetic ferric oxide beads. The latter were most effectively incorporated by the cells irrespective of the bead size (50 nm or 1 µm diameter) and irrespective of the mode of peptide conjugation to the beads via immobilized streptavidine, anti-MUC1 antibody (C595, B27.29) or Helix pomatia agglutinin (in case of GalNAc substituted glycopeptides). While more than 80% of mouse DCs were loaded with antigen-coated beads (1 µm) after 2h incubation time, a proportion of only 15 to 20% was reached with peptide-antibody complexes and below 5% with free peptide antigen. In accord with this, all attempts to identify peptide processing products after pulsing with free antigen or antigen-antibody complexes were unsuccessful.

Mouse DCs were adapted to AIMV medium and pulsed for 4h with a mixture of biotinylated glycopeptides (H1 – H6) conjugated to antibody- and lectin-coated beads. Three fractions, the magnetically separated cell pellet, the centrifuged cell pellet and the supernatant were analysed for the presence of proteolytic fragments derived from the 21mer GalNAc-peptides (Fig. 1). Peptide isolation was performed by magnetic separation of streptavidine-coated beads after incubation with cell extracts or supernatant or by reversed-phase solid-phase extraction on Poros C18. The fractions were analysed by MALDI(tof) mass spectrometry to obtain the mass pattern of the peptide products and by nanoflow LC-ESI mass spectrometry in the MS/MS mode to get sequence information. The results presented in Fig. 1 indicate extensive fragmentation of the glycopeptides, the supernatant containing mostly processed products, while the cellular fractions contained also considerable proportions of residual 21mer glycopeptides. The AHG21 glycopeptides AHGVTSAPD(E)T(S)RPAPGS TAPP(A)A (substituted with one GalNAc residue) were identified at m/z 2249.0 and 2223.0, respectively, corresponding to the masses of N-thiopropionylated H1 to H3 (m/z 2249.0) and H4 to H6 (m/z 2223.0). The only products identified were registered at m/z 1695.7 (P1) and m/z 1669.7 (P2), respectively, corresponding to the GalNAc containing peptide fragments SAP16. The sequence of the two peptide products (P1, P2) were confirmed by MS/MS on a Qtof2 instrument to comprize 16 aa long C-terminal portions of the AHG21 glycopeptides (Fig. 2),

P1 SAPDTRPAPGSTAPPA (SEQ ID NO 7) and

P2 SAPESRPAPGSTAPAA, (SEQ ID NO 8)

both containing GalNAc at Thr / Ser10 or Thr17 (numbering according to the AHG21 sequence). No SAP16 peptides devoid of GalNAc were registered at m/z 1492 and 1466, respectively, indicating that proteolysis of AHG21 with GalNAc at Thr5 adjacent to the cleavage site had not occurred. Sequencing of the remaining AHG21 glycopeptides by MS/MS analysis revealed that the GalNAc was preferentially located at Thr5, which is in accord with the assumption that undigested protease-resistant AHG21 glycopeptides were identical to H1 and H4, respectively (Fig. 2). The five aa N-terminal proteolytic fragment AHGVT (N-thiopropionylated) was not detected in any of the spectra. Control experiments with DC primed AIMV media without antigen (4h) indicated that no proteolytic activity had been secreted into the medium resulting in detectable endopeptidase cleavage of TAP25 peptide after 24h incubation at 37 °C. However, after adjustment of the supernatant to

conditions optimal for cystein proteases (pH5.5, 1 mM dithiotreitol) minor exopeptidase cleavage of the peptide was registered in the mass spectrum (data not shown). Hence the SAP16 fragments detected in the supernatants of antigen-pulsed cells can be regarded as cellular products and not as extracellular products of secreted proteases.

In vitro proteolysis of MUC1 glycopeptides with human cathepsin L

To verify the processing data obtained with mouse DCs and to confirm the proposed identity of the preferentially involved protease(s) we performed a series of *in vitro* digestions with cathepsin L and selected peptide substrates (Tab. 1, Fig. 3). Using standard conditions for cystein proteases and incubation times of 3h the enzyme was able to cleave all non-glycosylated MUC1 repeat peptides quantitatively, except for the variant GST20-AES peptide (80% cleavage). Irrespective of the starting amino acid position in the repeat sequence (TAP, AHG, GST) and of the length of the peptides (20mer, 21mer, 25mer, 100mer), cathepsin L cleaved specifically between Thr-Ser in the VTSA motif of the repeat peptide. Besides this preferential cleavage site, which is in accord with the *in vivo* data, minor activities of the enzyme preparation were found to be directed to the adjacent positions Val-Thr (TSA17 at m/z 1959) and Ser-Ala (APD15 at m/z 1771). To exclude the possibility that aminopeptidases could be responsible for the generation of these minor products, a protected substrate carrying a biotin label at the amino terminus was used as substrate (Fig. 3A). Also in agreement with DC-mediated processing was the finding that O-glycosylated peptides, carrying GalNAc or Galβ1-3 GalNAc, were effectively digested (Tab. 1, Fig. 3B). Moreover, the position of glycan attachment to one of the tree threonines (Thr5, Thr10, Thr17) in the AHG21 sequence was found to be critical as suggested by the *in vivo* data. While GalNAc or Galβ1-3GalNAc in positions more distant from the cleavage site (Thr10, Thr17) had no influence on the cleavage by cathepsin L, the glycopeptides H1 and H4, both being glycosylated at the cleavage site (Thr5), were stable to proteolysis. Minor exopeptidase activity was detectable (in case of these glycopeptides) in the cathepsin L preparation from human liver. As a control, human cathepsin D was tested with a selected panel of MUC1 repeat peptides and glycopeptides and found to be unable to use any of these as a substrate, even if incubation times of up to 24h were chosen. It can be concluded that proteolytic activity in the human cathepsin L preparation recapitulated all major aspects of MUC1 glycopeptide processing in vivo by mouse DCs.

In vitro proteolysis of MUC1 glycopeptides with enzymes in low-density endosomal fractions from mouse dendritic cells

Mouse dendritic cells were ruptured in the absence of protease inhibitors and the supernatant after removal of nuclei was centrifuged in a Percoll gradient. The gradient fractions were tested for proteolytic activity using TAP25 as a substrate and incubation conditions optimized for cystein proteases (Fig. 4). Low density endosomes were separated from lysosomes according to the registration of marker proteins (β-hexosaminidase) and demonstrated to contain a cystein protease inhibitable with Z-Leu-Leu-Leu-fluoromethyl ketone and with a site-specificity related to human cathepsin L. Only enzymes in fractions with a density around 1.037 g/ml cleaved TAP25 peptide and H3/H6 glycopeptides at Thr-Ser yielding SAP16, while all fractions, in particular those with densities above 1.054 g/ml, contained considerable activities of carboxy peptidase(s) (Fig. 4).

The presented work reveals insight for the first time into the processing of the human glycoprotein tumor antigen MUC1 by DCs. Using state-of-the-art methodologies for the structural characterization of peptides/glycopeptides this study was able to answer four important questions regarding MUC1 proteolysis by APCsin the MHC class II pathway: 1) Which are the cleavage sites in the MUC1 repeat peptide ? 2) In which way do O-linked glycans affect proteolytic cleavage ? 3) Are core-type glycans removed prior to proteolytic processing ? 4) Which of the enzymes involved in the processing machinery are responsible for proteolytic cleavage of MUC1 repeat peptides ? Our results suggest that MUC1 repeats are cleaved only at one site, between Thr-Ser in the VTSA motif. During *in vivo* processing the core-type glycans GalNAc and Galβ1-3-GalNAc were not removed, but inhibited the cleavage, if they were located adjacent to the cleavage site. The above aspects, in particular the site specificity of cleavage and the site-dependent effects of carbohydrates were perfectly simulated *in vitro* using human cathepsin L. Because this enzyme belongs to the cysteine protease family related to papain, and is claimed to be involved in antigen processing (7,8), we further confirmed the involvement of cathepsin L (or a closely related enzyme species) in MUC1 repeat proteolysis by specific inhibition. It can be anticipated that antigen processing in late endosomes is mediated by a family of proteases with partially overlapping, but still distinct specificities. Hence, the *in vitro* data on cathepsin L cleavage of MUC1 presented in this paper do not exclude the possible involvement of other, cathepsin L related enzymes in the processing machinery and in the specific cleavage of MUC1.

The peptide SAP20 or the glycosylated derivatives carrying O-linked glycans at the more C-terminal Thr/Ser positions may represent the preferential, if not exclusive processing products of MUC1 repeats in DCs. The structural features of these products match findings from an immunological study performed in parallel[1]. According to this work, glycans remain intact during processing of MUC1 glycopeptides by DCs, but influence activation of T cell hybridoma clones in a site-specific manner. Clone VF5, reactive to a peptide epitope that comprizes the DTR motif, was activated by DCs pulsed with AHG21 glycopeptides which carried glycans at Ser16 or Thr17. No increase in IL-2 production by this clone was measurable, however, if the glycans were located at the proposed epitope or at the Thr/Ser positions adjacent to the cleavage site defined in the present study (Thr5, Ser6). Hence the O-linked glycans can alter site-specifically proteolytic processing or presentation of the MHC class II-restricted glycopeptide. That glycans negatively affect presentation of glycopeptides on MHC class II can not be generalized, since loading experiments demonstrated effective MHC class II presentation of glycopeptides and activation of T cells (16). There is evidence that glycans have not only negative effects in the DC-mediated processing and T cell activation. Consistent with this, we have generated and characterized a T cell hybridoma VF9 that is reactive to MHC class II-presented MUC1 glycopeptides which carriy Galβ1-3GalNAc at Thr5 in the AHG21 sequence[1]. The specific epitope of this T cell hybridoma clone VF9 implies that a 21meric glycopeptide passes the processing machinery uncleaved and is presented as such on MHC class II proteins. Thus, O-linked glycans not only survive the processing machinery of APCs and are presented on MHC class II proteins they can be also involved in the epitope structure recognized by T cell receptors. Based on this it can be hypothetically generalized that glycans may contribute in this way to a greater diversification of the immune response to the redundant protein core of MUC1.

Tumor-associated MUC1, in particular the glycoforms from breast cancer cells, have been claimed to exhibit underglycosylated protein cores (17), referring to both, to truncated chain lengths and to a reduced number of glycosylated sites per repeat peptide. We could recently show that this assumption is only partially correct, since the structural analysis of MUC1 samples that were recombinantly expressed in four different breast cancer cell lines revealed increased O-glycosylation densities (18). Moreover, the patterns of O-linked chains differed strongly between individual cell lines indicating that there is no common breast cancer-

associated profile. Finally, the proposed shift from core2-based glycans to core1 was not verified, on the contrary, core2 was demonstrated to form the prevalent structural basis of the cancer-associated glycans. An average profile of O-linked glycans determined for MUC1 from pooled ascites samples of breast and pancreatic cancer patients was also in accord with a more complex, core2-based glycosylation (19). Interestingly, this glycoform of the mucin represents a weak immunogen in the MHC class I pathway (20) and is non-immunogenic in the MHC class II pathway (14). The latter phenomenon has been assigned to an entrapment in early endosomes of DCs mediated by multivalent, high-avidity interaction with the mannose receptor (14). It can be concluded, accordingly, that O-glycosylation of MUC1 interferes first of all with trafficking of endocytosed mucin. Later on, if late endosomal compartments are accessible for the antigen, other modes of interference mediated by O-linked glycans could also come into play, like a sterical inhibition of proteolysis. The existence of such restrictions introduced by site-specific O-glycosylation became evident in the present study, since glycans linked to Thr / Ser in the VTSA motif of MUC1 repeats prevented processing of the glycopeptides. The specific proteolysis of the MUC1 peptide by cathepsin L would explain also the observation that a Tn-100mer peptide with a total of 15 GalNAc residues linked to each Thr5, Ser16 and Thr17 (numbered according to the AHG20 sequence) was unable to activate Th cell clones [1]. Hence the site-specificity of glycan substitution, not necessarily the glycan structure, has to be considered in the design of cancer vaccines. Because not all glycosylated epitopes become available for MHC class II presentation, SAP20 and its glycosylated derivatives represent a "pre-processed" form suitable for internal or external loading on MHC class II molecules in immunotherapeutic approaches. In loading experiments with SAP20 the glycosylation-dependent effects on the binding to MHC class II proteins and to the T cell receptors can be studied by systematic variation of the substitution sites and structures of the glycans.

References

1. Kotera, Y., Fontenot, D.J., Piecher, G., Metzgar, R.S., Finn, O.J. (1994) *Cancer Res.* **54**, 2856 - 2860
2. Barnd, D.L., Lan, M.S., Metzgar, R.S., Finn, O.J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7159 - 7163
3. Hilkens, J., Kroezen, V., Bonfrer, J.M., De Jong-Bakker, M., Bruning, P.F. (1986) *Cancer Res.* **46**, 2582 - 2587
4. Agrawal, B., Reddish, M.A., Krantz, M.J., Longenecker, M.B. (1995) *Cancer Res.* **55**, 2257 - 2261
5. Karanikas, V., Hwang, L.A., Pearson, J., Ong, C.S., Apostolopoulos, V., Vaughan, H., Xing, P.X., Jamieson, G., Pietersz, G., Tait, B., Broadbent, R., Thynne, G., McKenzie, I.F.C. (1997) *J. Clin. Invest.* **100**, 2783 - 2792
6. Goydos, J.S., Elder, E., Whiteside, T.L., Finn, O.J., Lotze, M.T. (1996) *J. Surg. Res.* **63**, 298 - 304
7. Nakagawa, T.Y., Rudensky, A.Y. (1999) *Immunol. Rev.* **172**, 121 - 129
8. Honey, K., Duff, M., Beers, C., Brissette, W.H., Elliott, E.A., Peters, C., Maric, M., Cresswell, P., Rudensky, A. (2001) *J. Biol. Chem.* **276**, 22573 - 22578
9. Shen, Z., Reznikoff, G., Dranoff, G., Rock, K.L. (1997) *J. Immunol.* **158**, 2723 - 2730
10. Karsten, U., Diotel, C., Klich, G., Paulsen, H., Goletz, S., Müller, S., Hanisch, F.-G. (1998) *Cancer Res.* **58**, 2541 - 2549
11. Hanisch, F.-G., Müller, S., Hassan, H., Clausen, H., Zachara, N., Gooley, A.A., Paulsen, H., Alving, K., Peter-Katalinic, J. (1999) *J. Biol. Chem.* **274**, 9946 - 9954
12. Hanisch, F.-G., Reis, C.A., Clausen, H., Paulsen, H. (2001) *Glycobiology* **11**, 731 - 740
13. Müller, S., Alving, K., Peter-Katalinic, J., Zachara, N., Gooley, A.A., Hanisch, F.-G. (1999) *J. Biol. Chem.* **274**, 18165 - 18172
14. Hiltbold, E.M., Vlad, A.M., Ciborowski, P., Watkins, S.C., Finn, O.J. (2000) *J. Immunol.* **165**, 3730 - 3741
15. Barnes, K.A., Mitchell, R.N. (1995) *J. Exp. Med.* **181**, 1715 – 1727
16. Jensen, T., Hansen, P., Galli-Stampino, L., Mouritsen, S., Frische, K., Meinjohanns, E., Meldal, M., Werdelin, O. (1997) *J. Immunol.* **158**, 3769 - 3778

17. Lloyd, K.L., Burchell, J., Kudryashov, V., Yin, B.W.T., Taylor-Papadimitriou, J. (1996) *J. Biol. Chem.* **271**, 33325 - 33334

18. Müller, S., Hanisch, F.-G. (2002) *J. Biol. Chem.* **277**, 26103 – 26112

19. Beatty, P., Hanisch, F.-G., Stolz, D.B., Finn, O.J., Ciborowski, P. (2001) *Clin. Cancer Res.* **7**, 781 – 787

20. Hiltbold, E.M., Alter, M.D., Ciborowski, P., Finn, O.J. (1999) *Cell. Immunol.* **194**, 143 – 149

Footnotes

1) Vlad, A., Müller, S., Cudic, M., Paulsen, H., Otvos, L., Hanisch, F.-G., Finn, O.J. (2002) *J. Exp. Med.*, manuscript submitted (unpublished).

In vitro proteolysis of MUC1 glycopeptides and peptides with human cathepsins L and D

A *Cathepsin L*

Substrate	Structure / Sequence	Average mass of product ion
H11	① AHGVT-SAPDTRPAPGSTAPPA	2162.8 (AHG21 + HexNAc)
H12	① AHGVT SAPDTRPAPGSTAPPA	1696.5 (SAP16 + HexNAc)
H13	① AHGVT SAPDTRPAPGSTAPPA	1696.7 (SAP16 + HexNAc)
100mer	(HGVT SAPDTRPAPGSTAPPA)$_5$ SAPDTRPAPGSTAPPAHGVT	1888.0 (-H_2O) (SAP20)
Tn-100mer	① ① ① (HGVT-SAPDTRPAPGSTAPPA)$_5$	- 12 - - -
TAP25	TAPPAHGVT SAPDTRPAPGSTAPPA	1491.8 (SAP16)
A3	② AHGVT SAPDTRPAPGSTAPPA	1856.9 (SAP16 + Hex-HexNAc)
GST20-AES	GSTAPAAHGVT SAPESRPAP	910.8 (SAP9)

Table 1A

In vitro proteolysis of MUC1 glycopeptides and peptides with human cathepsins L and D

B *Cathepsin D*

Substrate	Structure / Sequence	Average mass of product ion
H11	AHGVT-SAPDTRPAPGSTAPPA (① on T)	2161.0 (AHG21 + HexNAc)
H12	AHGVT-SAPDTRPAPGSTAPPA (① on T)	2161.0 (AHG21 + HexNAc)
H13	AHGVT-SAPDTRPAPGSTAPPA (① on T)	2161.0 (AHG21 + HexNAc)
100mer	$(HGVT\text{-}SAPDTRPAPGSTAPPA)_5$	- 9 - -

Table 1B

The substrates (10 µg in 20 µl 0.1M sodium acetate buffer, pH 5.5, containing 1 mM EDTA) were incubated with cathepsin D for 24h or with cathepsin L for 3h (in the presence of 1 mM DTT) at 37°C. ① refers to O-linked GalNAc, ② to O-linked Galβ1-3GalNAc.

Immunogenic MUC1 glycopeptides

PATENT CLAIMS

1. A peptide fragment of the tandem repeat domain of MUC1, starting with the amino acid sequence SAP and comprising at least 9 amino acids.

2. The MUC1 peptide fragment of claim 1, wherein the MUC1 fragment is in the range of 10 to 25 amino acids.

3. A MUC1 peptide fragment which comprises an amino acid of SEQ ID NO 1-4, or variants thereof, wherein said variants comprise one or more insertions, substitutions and/or deletions as compared to the sequence of SEQ ID No. 1-4, and wherein the biological activity is substantially equal to the activity of the peptide comprising the unmodified amino acid sequence of SEQ ID No. 1-4, and provided that said variants start with the amino acid sequence SAP and comprise at least 9 amino acids.

4. The MUC1 peptide fragment of one or more of claims 1-3, wherein one or more of the threonines or serines of the peptide fragment are O-glycosylated.

5. The MUC1 peptide fragment of claim 4, which corresponds to one of SEQ ID No. 1-4, wherein the amino acid is glycosylated at Thr 5 and/or 12.

6. A nucleic acid, which encodes a MUC1 peptide fragment of any of claim 1-3.

7. A method of producing the MUC1 peptide fragments of one or more of claims 1-5, comprising the following steps:

a) providing a peptide comprising the tandem repeat domain of MUC1 or a part thereof, which part at least contains one repeating unit of said tandem repeat domain of MUC1,

b) contacting the peptide of a) with an effective amount of cathepsin-L or a closely related enzyme hereof, thereby cleaving the peptide, and

c) isolating the fragments produced in b).

8. The method of claim 7, wherein the peptide provided in step a) is natural MUC1 derived from human milk fat membranes, from human tumor ascites or from human breast carcinoma cell lines or is represented by SEQ ID NO 5, 6, 9 or 10.

9. The method of claim 7 or 8, wherein one or more of the amino acids of the peptide provided in step a) is O-glycosylated, provided that the peptide is not glycosylated at the cleaving site of cathepsin-L.

10. The method of one or more of claims 7 – 9, wherein one or more of the threonines or serines of the peptide fragment isolated in c) are O-glycosylated.

11. A MUC1 peptide fragment, which is obtainable by a method of one or more of claims 7-10.

12. An ex vivo-method of producing a population of autologous antigen presenting cells (APC's), which are capable of inducing effective immune responses against MUC1, comprising the steps of

a) providing autologous APC's from a tumor patient,

a) contacting the autologous APC's from the tumor patient with an effective amount of a MUC1 peptide fragment of one or more of claims 1-5 or claim 11 under conditions which allow endocytosis, processing and MHC class II presentation of the peptide fragments by said APC's, and

b) isolating said MUC1 peptide presenting APC's for the purpose of immunotherapeutic application in the patient.

13. The method of claim 12, wherein the MUC1 peptide fragments in b) are bound to ferric oxide beads.

14. An ex vivo-method of producing genetically engineered APC's, which are capable of inducing effective immune responses against MUC1, comprising the steps of

a) providing a nucleic acid, which codes for one of the MUC1 peptide fragments of claim 1-5 or 11,

b) transfecting the APC's with said nucleic acid, and

c) selecting APC's, which present said MUC1 peptide fragments in an MHC II restricted manner.

15. The method of claim 14, wherein the nucleic acid is provided in an expression vector in step a).

16. An APC, which is obtainable by the method of one or more of claims 12-15.

17. The APC of claim 16, which is a dendritic cell or a B cell.

18. A therapeutic composition, comprising a therapeutically effective amount of the MUC 1 peptide fragment of one or more of claims 1-5 or claim 11 or the APC's of one or more of claims 16 or 17 and a pharmaceutically acceptable carrier.

19. The therapeutic composition of claim 18, which is a vaccine.

20. Use of the therapeutic composition of claim 18 or 19 in the treatment of breast cancer or other MUC1-positive carcinomas including colorectal, pancreatic and gastric carcinomas.

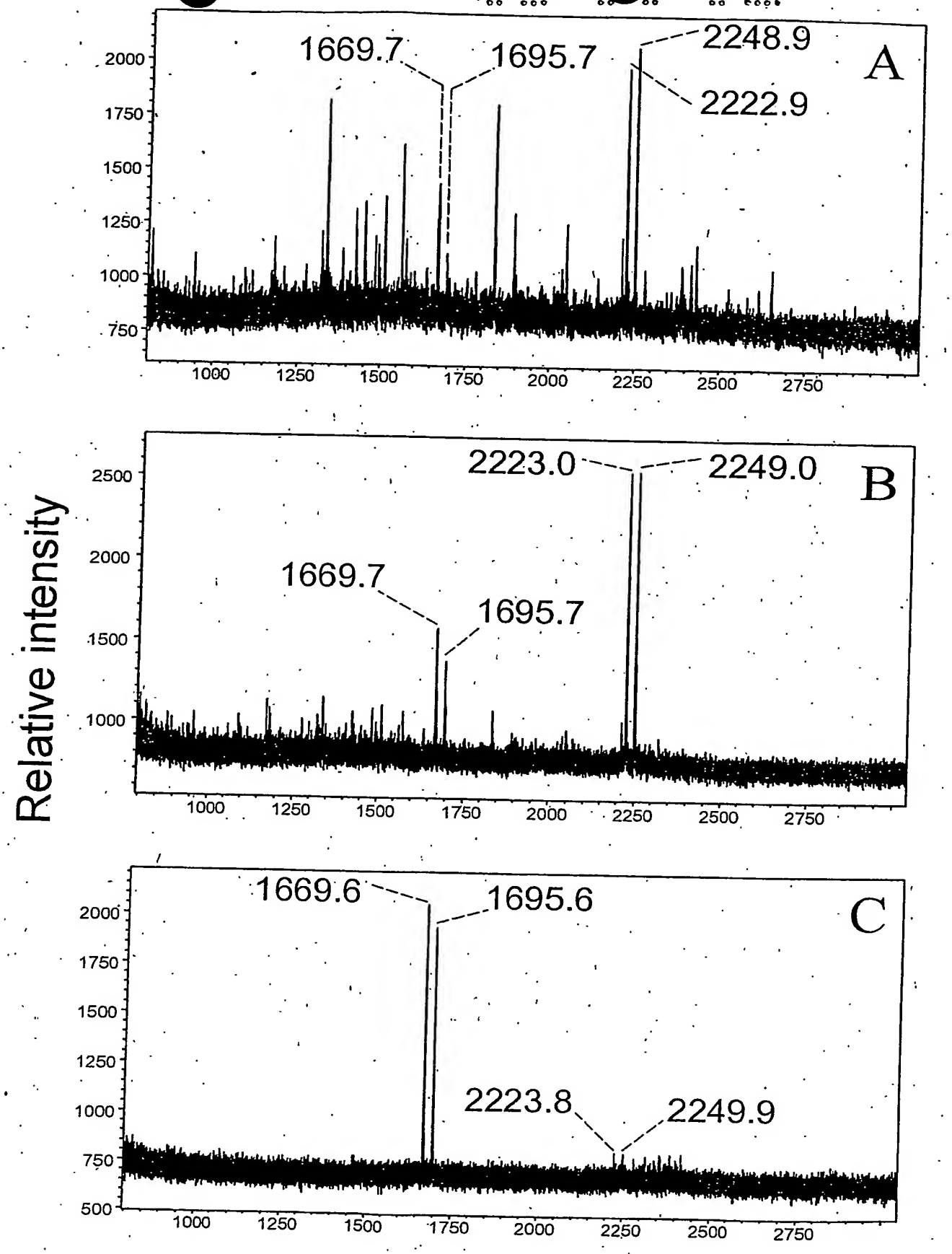
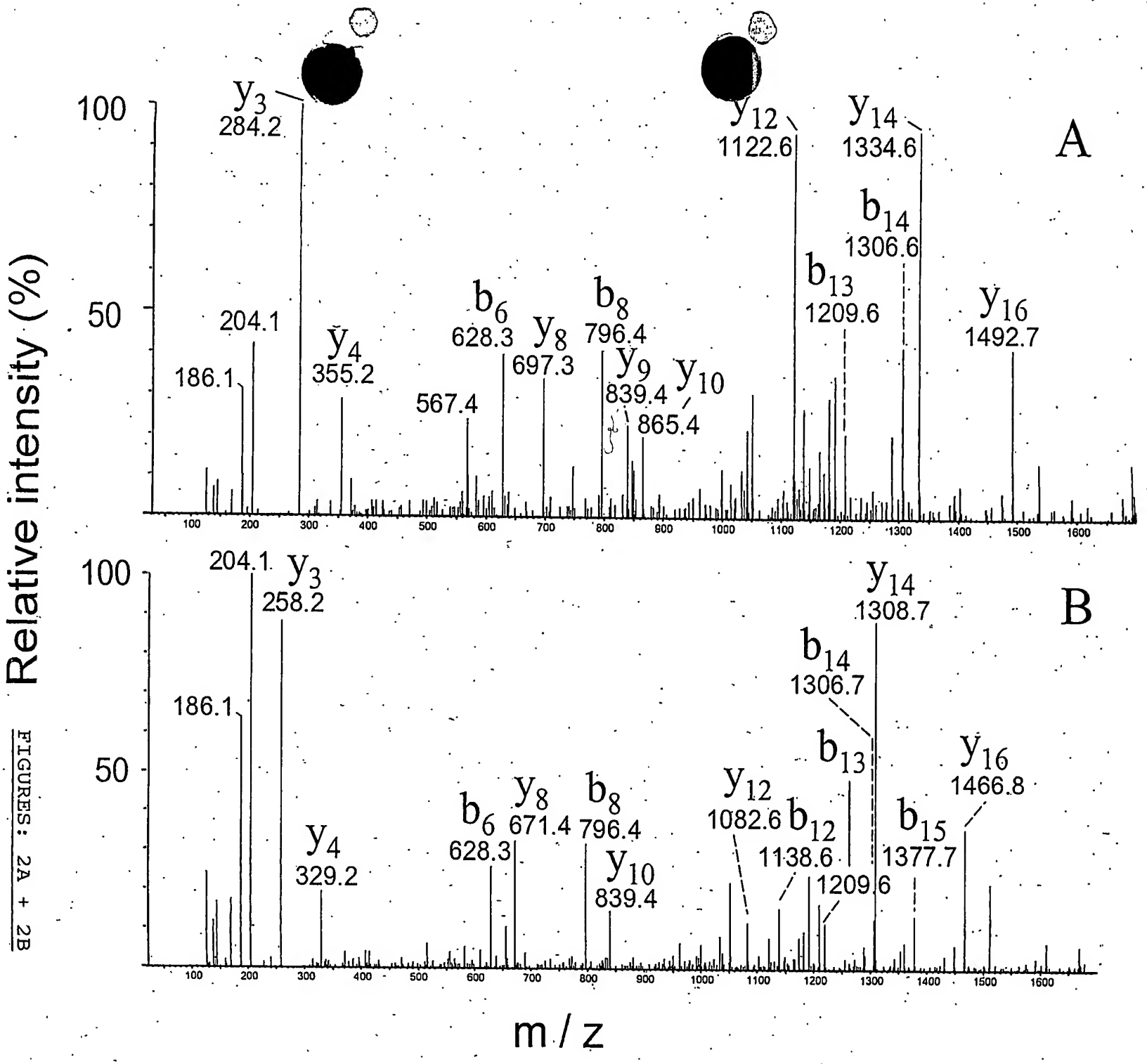

1669.7
1695.7
2248.9
2222.9
A
2000
1750
1500
1250
1000
750
1000 1250 1500 1750 2000 2250 2500 2750
2223.0
2249.0
B
1669.7
1695.7
2500
2000
1500
1000
1000 1250 1500 1750 2000 2250 2500 2750
1669.6
1695.6
C
2223.8
2249.9
2000
1750
1500
1250
1000
750
500
1000 1250 1500 1750 2000 2250 2500 2750
Relative intensity
m / z

FIGURE: 1



FIGURES: 2A + 2B

C

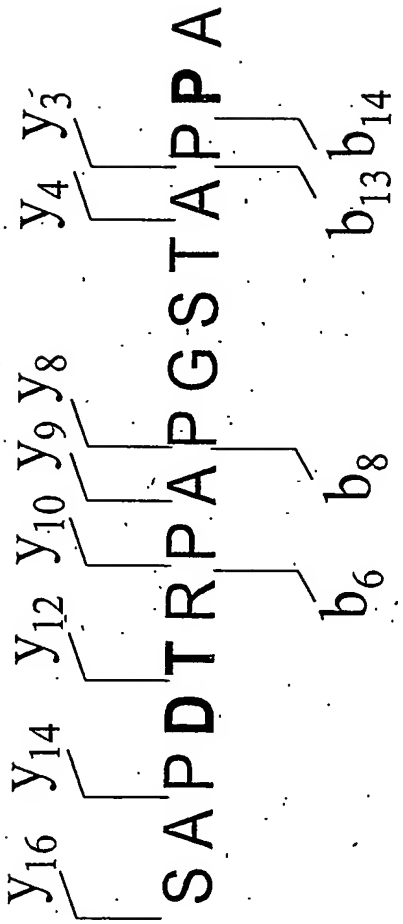

SAPDTRPAPGSTAPPA
y16 y14 y12 y10 y9 y8 y4 y3
b6 b8 b13 b14
SAPESRPAPGSTAPAA
y16 y14 y12 y10 y8 y4 y3
b6 b8 b12 b13 b15

FIGURE: 2C

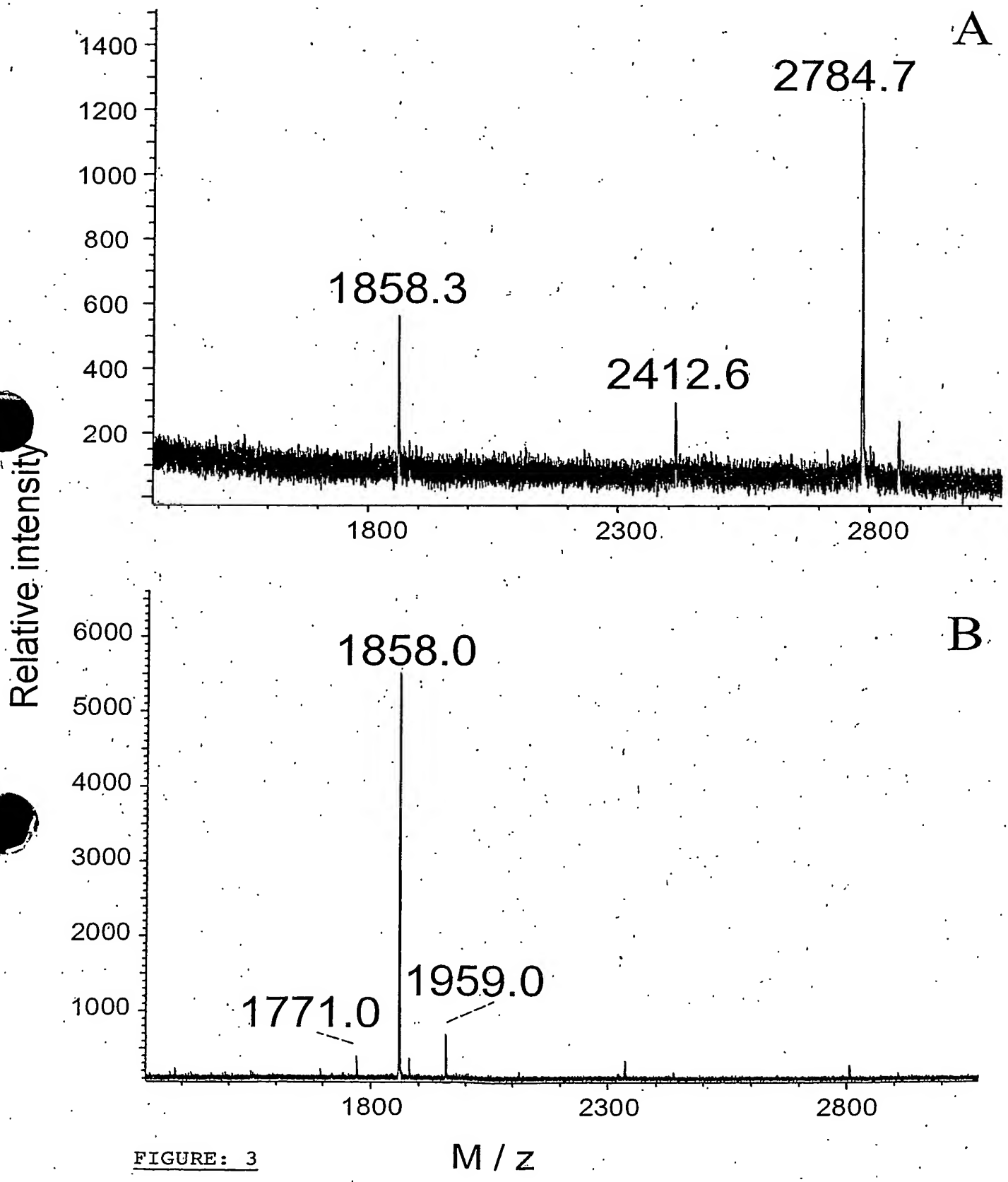

A
2784.7
1858.3
2412.6
1400
1200
1000
800
600
400
200
1800
2300
2800
B
1858.0
1959.0
1771.0
6000
5000
4000
3000
2000
1000
1800
2300
2800
Relative intensity
M / z

FIGURE: 3

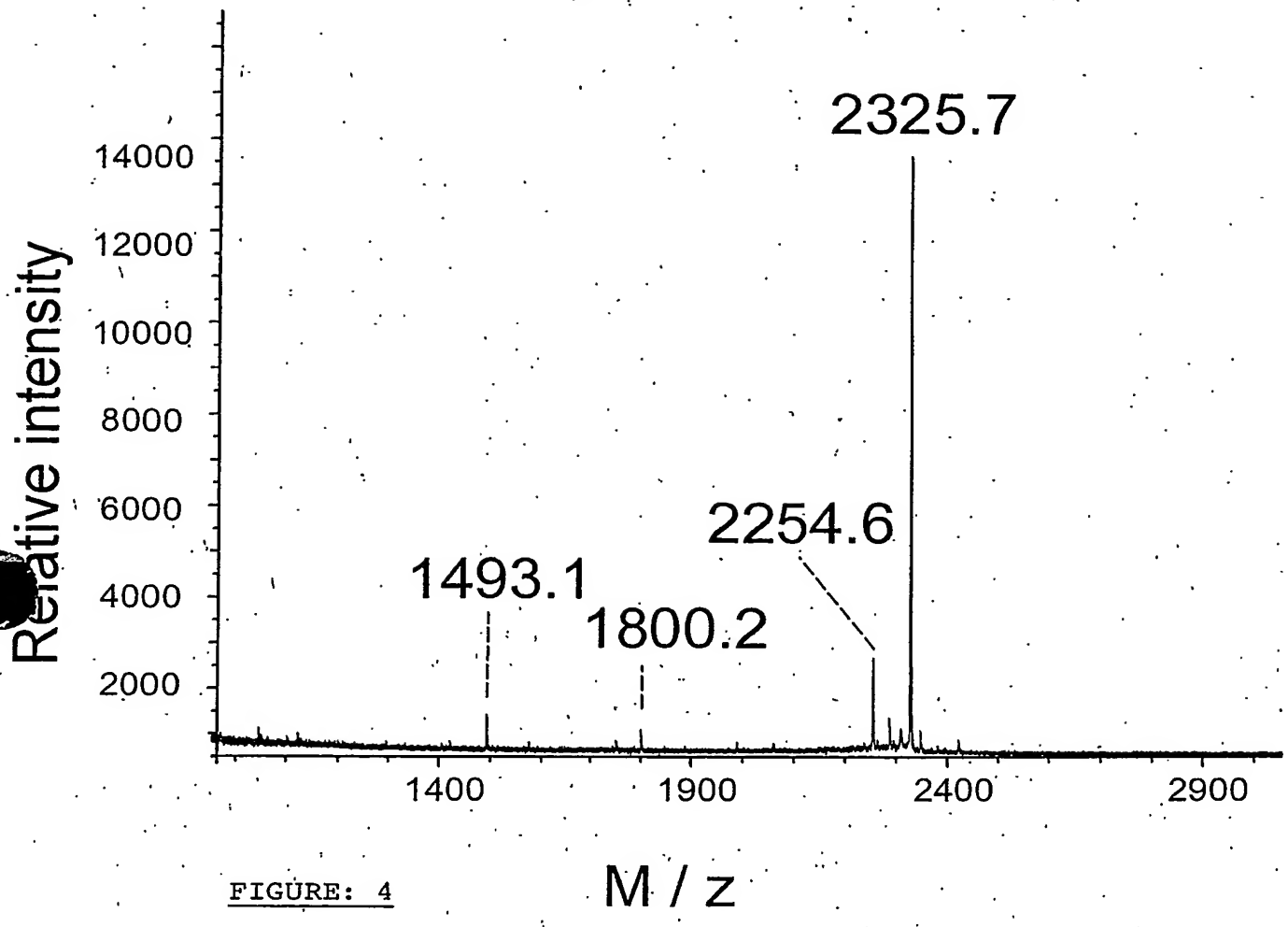

2325.7
14000
12000
10000
8000
6000
4000
2000
Relative intensity
2254.6
1493.1
1800.2
1400
1900
2400
2900
M / z

FIGURE: 4

SEQUENCE LISTING

<110> Cell Center Cologne GmbH

<120> Immunogenic MUC1 glycopeptides

<130> P14655

<160> 10

<170> PatentIn version 3.1

<210> 1
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> SAP20

<400> 1

Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala
1 5 10 15

His Gly Val Thr
20

<210> 2
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> SAP20

<400> 2

Ser Ala Pro Glu Ser Arg Pro Ala Pro Gly Ser Thr Ala Pro Ala Ala
1 5 10 15

His Gly Val Thr
20

<210> 3
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> SAP20

<400> 3

Ser Ala Pro Glu Ser Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala
1 5 10 15

His Gly Val Thr

<210> 4
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> SAP20

<400> 4

Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Ala Ala
1 5 10 15

His Gly Val Thr
20

<210> 5
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> H1-H3

<400> 5

Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser
1 5 10 15

Thr Ala Pro Pro Ala
20

<210> 6
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> H4-H6

<400> 6

Ala His Gly Val Thr Ser Ala Pro Glu Ser Arg Pro Ala Pro Gly Ser
1 5 10 15

Thr Ala Pro Ala Ala
20

<210> 7
<211> 16
<212> PRT
<213> Artificial Sequence

<220>

<223> P1

<400> 7

Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala
1 5 10 15

<210> 8
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> P2

<400> 8

Ser Ala Pro Glu Ser Arg Pro Ala Pro Gly Ser Thr Ala Pro Ala Ala
1 5 10 15

<210> 9
<211> 25
<212> PRT
<213> Artificial sequence

<220>
<223> TAP25

<400> 9

Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg Pro
1 5 10 15

Ala Pro Gly Ser Thr Ala Pro Pro Ala
20 25

<210> 10
<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> GST20-AES

<400> 10

Gly Ser Thr Ala Pro Ala Ala His Gly Val Thr Ser Ala Pro Glu Ser
1 5 10 15

Arg Pro Ala Pro
20